

Claims

1 1. A method of docking a ligand to a protein
2 comprising:

3 performing a pre-docking conformational search to
4 generate multiple solution conformations of the ligand;

5 generating a binding site image of the protein,
6 said binding site image comprising multiple hot spots;

7 matching hot spots of the binding site image to
8 atoms in at least one solution conformation of the
9 multiple solution conformations of the ligand to obtain
10 at least one ligand position relative to the protein in
11 a ligand-protein complex formation; and

12 optimizing the at least one ligand position while
13 allowing translation, orientation and rotatable bonds
14 of the ligand to vary, and while holding the protein
15 fixed.

Sub
C1
1 2. The method of claim 1, wherein said performing the
2 pre-docking conformational search comprises creating a
3 database of the multiple solution conformations and storing
4 said three-dimensional database for subsequent use by said
5 matching.

1 3. The method of claim 2, wherein said database of
2 the multiple solution conformations comprises a
3 conformational database of a combinatorial library.

SUB
A4

1 4. The method of claim 1, wherein said performing the
2 pre-docking conformational search comprises:

3 randomly generating a plurality of uniformly
4 distributed conformations of the ligand;

5 minimizing a strain of each conformation of the
6 plurality of uniformly distributed conformations;

7 using the strain and a solvent accessible surface
8 area of each conformation to rank the conformations;
9 and

10 clustering the conformations and retaining a
11 desired number of top clusters of conformations, said
12 retained number of top clusters of conformations
13 comprising said multiple solution conformations of the
14 ligand.

1 5. The method of claim 1, wherein said generating the
2 binding site image includes at least one of creating a list
3 of apolar hot spots identifying points in the binding site
4 that are favorable for an apolar atom to bind, and
5 generating a list of polar hot spots identifying points in
6 the binding site that are favorable for a hydrogen bond
7 donor or acceptor to bind.

1 6. The method of claim 5, wherein said generating the
2 binding site image further comprises:

SUB
A3

3 placing a grid around the binding site of the
4 protein;

5 determining a hot spot search volume using said
6 grid;

7 determining hot spots using a grid-like search of
8 the hot spot search volume; and

9 for each type of hot spot, clustering the hot
10 spots and retaining a desired number of clusters of hot
11 spots with best scores, said desired number of clusters
12 comprising said multiple hot spots to be employed by
13 said matching.

1 7. The method of claim 1, wherein said matching
2 comprises:

3 matching atoms of the at least one solution
4 conformation to appropriate hot spots of the protein by
5 positioning the at least one solution conformation as a
6 rigid body into the binding site image;

7 defining a match, said match determining a unique
8 rigid body transformation; and

9 using the unique rigid body transformation to
10 place the at least one solution conformation of the
11 ligand into the binding site of the protein.

8. The method of claim 7, wherein said determining
the unique rigid body transformation comprises determining
the unique rigid body transformation that minimizes:

4

$$I(R, T) = \sum_{j=1}^3 |H_j - RA_j - T|^2$$

5

where:

6

H_j = a j^{th} hot spot of the protein;

7

A_j = a j^{th} atom of the at least one solution
conformation;

8

9

R = a 3×3 rotation matrix; and

10

T = a translation vector.

1

9. The method of claim 1, wherein said optimizing
comprises optimizing multiple protein-ligand complex
formations, said optimizing comprising:

2

3

4

eliminating each ligand position having a
predetermined percentage of ligand atoms with a steric
clash;

5

6

7

ranking remaining ligand positions using an atom
pairwise score with a desired atom score cutoff;

8

9

after ranking, clustering the ligand positions and
selecting a top number n of ligand positions; and

10

11

optimizing each ligand position of the n
positions, allowing the translation, rotation and
rotatable bonds of the ligand to vary.

12

13

1 11. A system for docking a ligand to a protein
2 comprising:

3 means for performing a pre-docking conformational
4 search to generate multiple solution conformations of
5 the ligand;

6 means for generating a binding site image of the
7 protein, said binding site image comprising multiple
8 hot spots;

9 means for matching hot spots of the binding site
10 image to atoms in at least one solution conformation of
11 the multiple solution conformations of the ligand to
12 obtain at least one ligand position relative to the
13 protein; and

14 means for optimizing the at least one ligand
15 position while allowing translation, orientation and
16 rotatable bonds of the ligand to vary, and while
17 holding the protein fixed.

Sub
C7 12. The system of claim 11, wherein said means for
performing the pre-docking conformational search comprises
means for creating a database of the multiple solution
conformations and for storing said three-dimensional
database for subsequent use by said matching.

13. The system of claim 12, wherein said database of
the multiple solution conformations comprises a
conformational database of a combinatorial library.

SUB
A7

1 14. The system of claim 11, wherein said means for
2 performing the pre-docking conformational search comprises:

3 means for randomly generating a plurality of
4 uniformly distributed conformations of the ligand;

5 means for minimizing a strain of each conformation
6 of the plurality of uniformly distributed
7 conformations;

8 means for using the strain and a solvent
9 accessible surface area of each conformation to rank
10 the conformations; and

11 means for clustering the conformations and
12 retaining a desired number of top clusters of
13 conformations, said retained number of top clusters of
14 conformations comprising said multiple solution
15 conformations of the ligand.

1 15. The system of claim 11, wherein said means for
2 generating the binding site image includes at least one of
3 means for creating a list of apolar hot spots identifying
4 points in the binding site that are favorable for an apolar
5 atom to bind, and means for generating a list of polar hot
6 spots identifying points in the binding site that are
7 favorable for a hydrogen bond donor or acceptor to bind.

SUB
A8

1 16. The system of claim 15, wherein said means for
2 generating the binding site image further comprises:

3 means for placing a grid around the binding site
4 of the protein;

5 means for determining a hot spot search volume
6 using said grid;

7 means for determining hot spots using a grid-like
8 search of the hot spot search volume; and

9 for each type of hot spot, means for clustering
10 the hot spots and for retaining a desired number of
11 clusters of hot spots with best scores, said desired
12 number of clusters comprising said multiple hot spots
13 to be employed by said matching.

17. The system of claim 11, wherein said means for
matching comprises:

3 means for matching atoms of the at least one
4 solution conformation to appropriate hot spots of the
5 protein by positioning the at least one solution
6 conformation as a rigid body into the binding site
7 image;

8 means for defining a match, said match determining
9 a unique rigid body transformation; and

10 means for using the unique rigid body
11 transformation to place the at least one solution
12 conformation of the ligand into the binding site of the
13 protein.

18. The system of claim 17, wherein said determining the unique rigid body transformation comprises determining the unique rigid body transformation that minimizes:

$$I(R, T) = \sum_{j=1}^3 |H_j - RA_j - T|^2$$

where:

H_j = a j^{th} hot spot of the protein;

A_j = a j^{th} atom of the at least one solution conformation;

R = a 3×3 rotation matrix; and

T = a translation vector.

19. The system of claim 11, wherein said means for optimizing comprises means for optimizing multiple protein-ligand complex formations, said means for optimizing comprising:

means for eliminating each ligand position having a predetermined percentage of ligand atoms with a steric clash;

means for ranking remaining ligand positions using an atom pairwise score with a desired atom score cutoff;

1 21. At least one program storage device readable by a
2 machine, tangibly embodying at least one program of
3 instructions executable by the machine to perform a method
4 of docking a ligand to a protein, comprising:

5 performing a pre-docking conformational search to
6 generate multiple solution conformations of the ligand;

7 generating a binding site image of the protein,
8 said binding site image comprising multiple hot spots;

9 matching hot spots of the binding site image to
10 atoms in at least one solution conformation of the
11 multiple solution conformations of the ligand to obtain
12 at least one ligand position relative to the protein;
13 and

14 optimizing the at least one ligand position while
15 allowing translation, orientation and rotatable bonds
16 of the ligand to vary, and while holding the protein
17 fixed.

Sub
CB 1 22. The at least one program storage device of claim
2 21, wherein said performing the pre-docking conformational
3 search comprises creating a database of the multiple
4 solution conformations and storing said three-dimensional
5 database for subsequent use by said matching.

1 23. The at least one program storage device of claim
2 22, wherein said database of the multiple solution
3 conformations comprises a conformational database of a
4 combinatorial library.

24. The at least one program storage device of claim 21, wherein said performing the pre-docking conformational search comprises:

randomly generating a plurality of uniformly distributed conformations of the ligand;

minimizing a strain and a solvent accessible surface area of each conformation of the plurality of uniformly distributed conformations;

using the strain of each conformation to rank the conformations; and

clustering the conformations and retaining a desired number of top clusters of conformations, said retained number of top clusters of conformations comprising said multiple solution conformations of the ligand.

25. The at least one program storage device of claim 21, wherein said generating the binding site image includes at least one of creating a list of apolar hot spots identifying points in the binding site that are favorable for an apolar atom to bind, and generating a list of polar hot spots identifying points in the binding site that are favorable for a hydrogen bond donor or acceptor to bind.

26. ~~The at least one program storage device of claim 25, wherein said generating the binding site image further comprises:~~

4 placing a grid around the binding site of the
5 protein;

6 determining a hot spot search volume using said
7 grid;

8 determining hot spots using a grid-like search of
9 the hot spot search volume; and

10 for each type of hot spot, clustering the hot
11 spots and retaining a desired number of clusters of hot
12 spots with best scores, said desired number of clusters
13 comprising said multiple hot spots to be employed by
14 said matching.

1 27. The at least one program storage device of claim
21, wherein said matching comprises:

2 matching atoms of the at least one solution
3 conformation to appropriate hot spots of the protein by
4 positioning the at least one solution conformation as a
5 rigid body into the binding site image;
6

7 defining a match, said match determining a unique
8 rigid body transformation; and

9 using the unique rigid body transformation to
10 place the at least one solution conformation of the
11 ligand into the binding site of the protein.

12 28. The at least one program storage device of claim
13 27, wherein said determining the unique rigid body
14

3 transformation comprises determining the unique rigid body
4 transformation that minimizes:

5
$$I(R, T) = \sum_{j=1}^3 |H_j - RA_j - T|^2$$

6 where:

7 H_j = a j^{th} hot spot of the protein;

8 A_j = a j^{th} atom of the at least one solution
9 conformation;

10 R = a 3x3 rotation matrix; and

11 T = a translation vector.

1 29. The at least one program storage device of claim
2 21, wherein said optimizing comprises optimizing multiple
3 protein-ligand complex formations, said optimizing
4 comprising:

5 eliminating each ligand position having a
6 predetermined percentage of ligand atoms with a steric
7 clash;

8 ranking remaining ligand positions using an atom
9 pairwise score with a desired atom score cutoff;

10 after ranking, clustering the ligand positions and
11 selecting a top number n of ligand positions; and

